(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 527 839 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent: **02.12.1998 Bulletin 1998/49**

(21) Application number: **91908963.1**

(22) Date of filing: **01.05.1991**

(51) Int Cl.[6]: **C12Q 1/70**, G01N 33/53, C12N 7/00, C12N 15/00, C12N 1/20, G01N 33/566, G01N 33/68

(86) International application number: **PCT/US91/02989**

(87) International publication number: **WO 91/17271 (14.11.1991 Gazette 1991/26)**

(54) **RECOMBINANT LIBRARY SCREENING METHODS**

SCREENING-VERFAHREN FÜR GENBANKEN (RECOMBINANT LIBRARIES)

PROCEDES DE TRIAGE DE BANQUES D'ADN RECOMBINANT

(84) Designated Contracting States:
AT BE CH DE DK ES FR GB GR IT LI LU NL SE

(30) Priority: **01.05.1990 US 517659**

(43) Date of publication of application: **24.02.1993 Bulletin 1993/08**

(60) Divisional application: **98200770.0 / 0 866 136**

(73) Proprietor: **AFFYMAX TECHNOLOGIES N.V.**
Willemstad, Curaçao (AN)

(72) Inventors:
- **DOWER, William, J.**
  Menlo Park, CA 94025 (US)
- **CWIRLA, Steven, E.**
  Palo Alto, CA 94303 (US)

(74) Representative: **Bizley, Richard Edward et al**
Hepworth, Lawrence, Bryer & Bizley
Merlin House
Falconry Court
Baker's Lane
Epping Essex CM16 5DQ (GB)

(56) References cited:
WO-A-91/18980

- **GENE. vol. 73, 1988, AMSTERDAM NL pages 305 - 318 S.F. PARMLEY ET AL.**
- **SCIENCE vol. 228, 14 June 1985, LANCASTER, PA US pages 1315 - 1317 G.P. SMITH**
- **Science, Volume 222, issued 18 Novembver 1983, R. A. YOUNG et al., "Yeast RNA Polymerase II Genes: Isolation with antibody Probes", pp. 778-782, see entire document**
- **Science, Volume 240, issued 20 May 1988, A. SKERRET et al., "Assembly of a Functional Immunoglobulin Fv Fragment in Escherichia coli", pages 1038-1041, see Abstract.**
- **Science, Volume 240, issued 20 may 1988, M. BETTER et al., "Escherichia coli Secretion of an Active Chimeric Antibody Fragment", pages 1041-1043, see Abstract.**
- **Nature, Volume 348, issued 06 December 1990, J. McCAFFERTY et al., "Phage antibodies: filamentous phage displaying antibody varible domains", pages 552- 554, see entire document.**
- **Science, Volume 246, issued 08 December 1989, W.D. HUSE et al., "Generation of a Large Combinat orial Library of the Immunoglobulin Repertoire in Phage Lamda", pages 1275-1281, see entire document.**

EP 0 527 839 B1

Description

Field of the Invention

The present invention relates generally to recombinant DNA technology and, more particularly, to methods for screening DNA libraries for DNA sequences that encode proteins of interest.

Background of the Invention

Isolating a gene which encodes a desired protein from a recombinant DNA library can be a daunting task. Hybridization probes may facilitate the process, but their use is generally dependent on knowing at least a portion of the sequence of the gene which encodes the protein. When the sequence is not known, DNA libraries have been expressed in an expression vector and antibodies have been used to screen plaques or colonies for the desired protein antigen. This procedure has been useful in screening small libraries, but sequences which are represented in less than about 1 in 10^5 clones are easily missed, and screening libraries larger than 10^6 can be difficult.

Antibody molecules are comprised of light and heavy polypeptide chains, each having a distinct variable (V) region, the combination of which produces an antigen binding region. Based on random combination events of heavy and light chains in any one antibody-producing cell, the potential repertoire of antibody heavy and light chain combinations may be as much as 10^{12} or greater. Thus, to sample a large fraction of this repertoire and obtain clones which express an antibody having a desired antigen binding specificity, an extremely large library may have to be constructed and screened.

Huse W.D. *et al* (1989) Science 146:1275-1281 reports on the preparation of cloning libraries of antibody heavy and light chains in phage lambda. The antibody chains assemble intracellularly inside the *E.coli* host independently of phage particle formation.

Parmley S.F. & Smith G.P. (1988) Gene 73:305-318 teaches the display and screening of single-chain peptides on fusion phages.

Methods are needed which facilitate the screening process, thereby enabling DNA sequences which encode proteins of interest, and particularly antibody molecules, to be more readily identified, recloned and expressed. Were such procedures available, it may become possible to probe an animal's entire antibody repertoire, for example, to obtain an antibody to a preselected target molecule. In this manner the difficulties and labor intensive process of generating monoclonal antibodies, regardless of the species of origin, by conventional hybridization or transformation of lymphoblastoid cells, may be avoided. Quite surprisingly, the present invention fulfills these and other related needs.

The present invention provides the use of a bacteriophage to display a multichain protein, wherein a first chain of the multichain protein is fused to a coat peptide on the outer surface of the bacteriophage, and a second chain of the multichain protein is complexed with the first chain.

The invention also provides a method for screening a DNA library for nucleotide sequences which encode a multichain protein comprising first and second polypeptide chains, which multichain protein binds specifically to a ligand, comprising:
effecting bacteriophage expression vector transformation of a host cell with:-

(i) a first nucleotide sequence member of the library that encodes the first chain fused to a sequence encoding a coat peptide of the bacteriophage; and

(ii) a second nucleotide sequence member of the library that encodes the second chain fused to a sequence encoding a signal peptide that directs periplasmic secretion of said second chain;

cultivating the transformed cell under conditions suitable for expression and assembly of bacteriophage particles and the multichain protein, wherein the multichain protein is displayed on the outer surface of the bacteriophage particles, optionally wherein expression of DNA library sequence members is inducible, induction of expression of the DNA library sequences preferably being delayed until assembly of at least one complete bacteriophage particle has occurred;

selecting bacteriophage particles encoding the multichain protein by means of the ligand and, if desired, further comprising the step of isolating the nucleotide sequences which encode the first and second chains of the multichain protein from the selected bacteriophage particles; optionally wherein the bacteriophage are harvested from the host cell culture before the selecting step.

The method therefore provides for screening a DNA library for a nucleotide sequence which encodes a protein of

interest. The method generally involves physically linking the protein of interest, in a biologically active form (usually having a binding activity), to the specific DNA sequence encoding that protein. This allows the isolation and identification of that DNA by means of affinity techniques relying on the binding activity of the protein of interest.

The second and any subsequent protein chain(s) are therefore expressed from the phage genome so as to be transported to the periplasm where they assemble with the first chain that is fused to the phage coat protein, which complex associates with the phage particle as it exits the cell. Phage particles which encode the protein of interest are selected by means of a ligand specific for the protein.

Particles which encode the protein of interest can be selected from the culture by an affinity enrichment technique. This is accomplished by means of a ligand specific for the protein of interest, such as an antigen if the protein of interest is an antibody. Repeating the affinity selection procedure provides an enrichment of clones encoding the desired sequences, which may then be isolated for sequencing, further cloning and/or expression.

The phage may be a filamentous phage, such as, for example, fd, fl, or M13. Typically the DNA library sequences are inserted in the 5' region of a gene which encodes a phage coat protein, such as pIII. Thus, a phage coat protein fused to a DNA library member-encoded protein is produced and assembled in to the viral particle.

The signal sequence may be an Omp A, pelB, phoA, pIII or β-lactamase signal sequence.

The transformed host cells may be lysed after cultivation and the bacteriophage particles may be selected from cellular debris; optionally the bacteriophage particles encoding the protein of interest may be enriched by repeating the selection step at least once.

A preferred multichain protein is an antibody or a binding fragment thereof, preferably a Fab fragment.

The said first nucleotide sequence member of the library may encode a protein which comprises an antibody heavy chain variable region, optionally said antibody heavy chain variable region being located at the amino terminus of said coat protein on said bacteriophage surface.

The first and second DNA library members preferably comprise amplified cDNA.

The method of the invention may further include the step of incorporating the multichain protein into a therapeutic, prophylactic or diagnostic composition.

Methods and compositions are therefore provided by the present invention for conveniently identifying clones of desired recombinantly-produced proteins. A method of affinity enrichment allows for the screening of libraries to identify clones having desired ligand specificities, where up to about 10^9 or more clones may be readily screened. This represents a significant improvement over the art, where conventional procedures typically allow about 10^6 clones in a DNA library to be screened, sometimes up to 10^7 clones but with proportionate increases in time and labor.

The invention permits linking the desired multichain protein, such as an antibody molecule, to the DNA which encodes it. By then enriching for the protein, such as by affinity techniques, for example, the DNA which encodes the protein is also enriched and may then be isolated. The DNA so obtained may then be cloned and expressed in other systems, yielding potentially large quantities of the desired protein, or may be subjected to sequencing and further cloning and genetic manipulations prior to expression.

The protein for which the DNA is enriched and cloned according to the present invention is typically an antibody or fragment thereof, but may also be any multichain protein which may be cloned from a nucleotide library. In addition to antibodies, such proteins may include, for example, growth hormones, interferons, interleukins, hormones, enzymes, zymogens, etc. Proteins which may be cloned are those for which specific binding partners (e.g. antigen or hapten when the desired protein is an antibody) have been identified.

When the protein of interest is an antibody of a desired binding specificity, the antibody may be of any of the known isotypes or subclasses for a particular species, and may be a two-chain binding complex or portion thereof. For instance, the variable antigen-binding regions of heavy (V_H) and light (V_L) chains may be identified and cloned; the binding fragments (F_v) or Fab encoded thereby may find use either as a binding fragment, joined to constant regions of heavy or light chains, or joined to other proteins having desired effector functions. The characteristics of the constant region domains will depend to a large extent on the use intended for the antibody, e.g., diagnostic and/or therapeutic applications, catalytic antibodies, etc.

Typically the protein's binding partner, e.g., an antigen or hapten when the protein is an antibody, is known, and the methods herein provide a means for creating and/or identifying a protein (and the DNA which encodes the protein) which specifically binds the binding partner of interest. Thus when the protein is an antibody the present invention provides a novel means for producing antibodies, particularly monoclonal antibodies, to predetermined antigens and antigenic determinants, thereby circumventing the laborious, time-consuming and often unpredictable process of conventional monoclonal antibody technology. Although murine monoclonal antibody production is often relatively straightforward, it is labor intensive. Furthermore, the development of human monoclonal antibodies by conventional approaches has been hampered by a variety of technical difficulties which, to a large extent, would be circumvented by the present invention.

DNA libraries are prepared from cells which are capable of encoding the desired protein. A variety of techniques exist for preparing a library, which may be prepared from either genomic DNA or cDNA. See, e.g., Sambrook et al.,

Molecular Cloning, A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989. It is understood that when referring herein to DNA it is meant to include both genomic and cDNA, unless otherwise specified. The cells which serve as the source of RNA or DNA may be any which are capable of encoding the protein of interest. Enrichment procedures and means for amplifying the regions containing the gene(s), if known, may be employed. For instance, when the desired protein is an antibody, RNA and cDNA may be prepared from spleen cells from unimmunized animals, from animals immunized with antigens or haptens of interest, hybridoma cells, or lymphoblastoid cells, for example. The use of spleen cells from unimmunized animals provides a better representation of the possible antibody repertoire, while spleen cells from immunized animals are enriched for sequences directed against epitopes of the immunizing antigen or haptens. The cells may be obtained from a variety of animal species, such as human, mouse, rat, lagomorpha, equine, bovine, avian, etc., the selection often dependent on the protein of interest and the use for which it is intended.

Amplification of sequences representing messenger RNA (mRNA) isolated from cells of interest, such as spleen or hybridoma cells, may be performed according to protocols outlined in, e.g., U.S. 4,683,202, Orlandi, et al. Proc. Natl. Acad. Sci. USA 86:3833-3837 (1989), Sastry et al., Proc. Natl. Acad. Sci. USA 86:5728-5732 (1989), and Huse et al. Science 246:1275-1281 (1989). Oligonucleotide primers useful in amplification protocols may be unique or degenerate or incorporate inosine at degenerate positions. Thus, for multi-chain immunoglobulins, primers would be generally used for amplification of sequences encoding the variable regions of both the heavy and light chains. Restriction endonuclease recognition sequences may be incorporated into the primers to allow for the cloning of the amplified fragment into a vector in a predetermined reading frame for expression.

Expression libraries containing the amplified cDNA are typically prepared in a vector such as a bacteriophage. The characteristics of the suitable bacteriophage depends on the specific embodiment employed, and will generally be those which conveniently allow insertion of the recombinant DNA into host cells by in vitro packaging or genetic transformation, which infect host cells capable of expressing the desired proteins, and whose DNA contains restriction sites, useful for cloning, which are located in regions of the phage genome where insertion of foreign DNA will not substantially disrupt essential functions of the phage.

As mentioned, the vectors may be engineered to efficiently clone the library amplification products. For example, oligonucleotides may be used to introduce the asymmetric restriction sites, a ribosome binding site at an optimal distance for expression of the cloned sequence, and cloning sites for the library amplification products.

To enrich for and isolate phage which contain cloned library sequences that encode a desired protein, and thus to ultimately isolate the nucleic acid sequences themselves, phage harvested from the bacterial debris are affinity purified. A ligand or binding partner specific for the desired cloned library protein is used in the affinity purification. For example, when the desired protein is an antibody which specifically binds a particular antigen or antigenic determinant, the antigen or determinant is used to retrieve phage having the desired protein on or a part of its outer surface. The ligand is typically adsorbed to an insoluble substrate, such as a particle or bead or plate. The phage so obtained may then be amplified by infecting into host cells. Additional rounds of affinity enrichment, as great as or greater than 10^4 fold per round, and amplification may be employed until the desired level of enrichment is reached or the target phage are no longer enriched relative to the background phage.

The enriched antibody-phage may also be screened with additional detection techniques such as expression plaque (or colony) lift (see, e.g., Young and Davis, Science, 222:778-782 (1983)), whereby the same or another binding partner is used as a probe. Screening may employ additional assays (for a catalytic activity, for example) which are used to detect, in situ, plaques expressing proteins having the desired characteristics. The phage obtained from the screening protocol are infected into cells, propagated, and the phage DNA isolated and sequenced, and/or recloned into a vector intended for gene expression in prokaryotes or eukaryotes to obtain larger amounts of the particular protein selected.

In accordance with the invention, the desired protein is transported to an extra-cytoplasmic compartment of the host cell, such as the bacterial periplasm, as a fusion protein with a viral coat protein. In this embodiment one chain of the desired multi-chain protein (such as an antibody) is expressed fused to a viral coat protein which is processed and transported to the cell inner membrane. Other chains are expressed with a secretion leader and thus are also transported to the periplasm or other intracellular but extra-cytoplasmic location. The chains (e.g., light and heavy chains) present in the extra-cytoplasm then assemble into a complete protein (or binding fragment thereof). The assembled molecules become incorporated into the phage by virtue of their attachment to the phage coat protein as the phage extrude through the host membrane and the coat proteins assemble around the phage DNA. The phage bearing the antibody complex may then be screened by affinity enrichment.

In this embodiment the synthesis and amplification of cDNAs is prepared as described above, and then is cloned into or near a vector sequence encoding a coat protein, where the vector is, or is derived from, a filamentous phage, such as fl, fd, Pfl, M13, etc. In a preferred embodiment the filamentous phage is fd-tet. The phage vector is chosen to contain a cloning site located in the 5' region of a gene encoding a phage coat protein, such as, for example, the pIII coat protein. An appropriate vector (e.g., fd-tet B1 which is described below) allows oriented cloning of foreign se-

quences so that they are expressed at or near the N-terminus of the mature coat protein.

A library is constructed by cloning the cDNA (e.g., the V_H region) from the donor cells into a coat protein gene (e.g., gene III, "gIII") cloning site. The cloned sequences of, for example, the V_H domains are ultimately expressed as polypeptides or proteins (of up to about 120 amino acids in the case of the V_H protein) fused to the N-terminus of the mature coat protein on the outer, accessible surface of the assembled phage particles. Although a large peptide fragment near the N-terminus of the coat protein may cause a decrease in the phage infectivity and/or yield compared to phage with much smaller, similarly-placed fragments, the larger fragments may still be effectively enriched by the procedures described herein.

Some peptides, because of their size and/or sequence, may cause severe defects in the infectivity of their carrier phage. This causes a loss of phage from the population during reinfection and amplification following each cycle of panning. To minimize problems associated with defective infectivity, DNA prepared from the eluted phage is transformed into host cells by electroporation or well known chemical means. The cells are cultivated for a period of time sufficient for marker expression, and selection is applied as typically done for DNA transformation. The colonies are amplified, and phage harvested as described below for a subsequent round(s) of panning.

The cDNA encoding the chain(s) not cloned into a phage coat protein may be cloned directly into an appropriate site (as described below) of the vector containing the first chain-coat protein library; or, preferably, the subsequent chain(s) may be cloned as a separate library in a different plasmid vector, amplified, and subsequently the fragments installed in the first chain-coat protein library vector. For example, when the first chain is an antibody heavy chain or binding fragment thereof, the ultimate destination of light chain V_L cDNA sequence is in a vector phage RF DNA that already contains a V_H sequence in a coat protein gene, thus randomly recombining V_H and V_L sequences in a single phage genome.

The second or subsequent chain of the desired multi-chain protein, such as V_L, is cloned so that it is expressed with a signal peptide leader sequence that will direct its secretion into the periplasm of the host cell. For example, several leader sequences have been shown to direct the secretion of antibody sequences in E. coli, such as OmpA (Hsiung, et al., Biotechnology 4:991-995 (1986)), pelB (Better, et al., Science 240:1041-1043 (1988)), phoA (Skerra and Pluckthun, Science 240:1038-1043 (1988)), and 8-lactamase (Zemel-Dreasen and Zamir, Gene 27:315-322 (1984)).

The cloning site for the subsequent chain cDNA's should be placed so that it does not substantially interfere with normal phage function. One such locus is the intergenic region as described by Zinder and Boeke, Gene 19:1-10 (1982). The V_L sequence is preferably expressed at an equal or higher level than the V_H/pIII product to maintain a sufficiently high V_L concentration in the periplasm to provide efficient assembly (association) of V_L with V_H chains.

Generally, the successful cloning strategy utilizing a phage coat protein, such as pIII of filamentous phage fd, will provide: (1) expression of a protein chain (or a first polypeptide chain when the desired protein is multi-chained, e.g., the V_H chain) fused to the N-terminus of a full sized (or nearly full sized) coat protein (e.g., pIII) and transport to the inner membrane of the host where the hydrophobic domain in the C-terminal region of the coat protein anchors the fusion protein in the membrane, with the N-terminus containing the chain protruding into the periplasmic space and available for interaction with a second or subsequent chain (e.g., V_L to form an F_v or Fab fragment) which is thus attached to the coat protein; (2) adequate expression of a second or subsequent polypeptide chain if present (e.g., V_L) and transport of this chain to the soluble compartment of the periplasm; and (3) will usually, but not necessarily, produce little or no interference with normal phage function or host cell viability.

The number of possible combinations of heavy and light chains probably exceeds 10^{12}. To sample as many combinations as possible depends, in part, on the ability to recover large numbers of transformants. For phage with plasmid-like forms (as filamentous phage), electrotransformation provides an efficiency comparable to that of phage λ transfection with in vitro packaging, in addition to a very high capacity for DNA input. This allows large amounts of vector DNA to be used to obtain very large numbers of transformants. The method described by Dower et al., Nucleic Acids Res., 16:6127-6145 (1988), may be used to transform fd-tet derived recombinants at the rate of about 10^7 transformants/μg of ligated vector into E. coli (such as strain MC1061), and libraries may be constructed in fd-tet B1 of up to about 3 x 10^8 members or more. Increasing DNA input and making modifications to the cloning protocol within the ability of the skilled artisan may produce increases of greater than about 10-fold in the recovery of transformants, providing libraries of up to 10^{10} or more recombinants.

The transformants are selected by growth in an appropriate antibiotic(s) which, in the case of the fd-tet vector, is tetracycline. This may be done on solid or in liquid growth medium. For growth on solid medium, the cells are grown at a high density (~10^8 to 10^9 tfs per m^2) on a large surface of, for example, L-agar containing the selective antibiotic to form essentially a confluent lawn. The cells and extruded phage are scraped from the surface and phage are prepared for first round of panning essentially as described by Parmley and Smith, Gene 73:305-318 (1988). For growth in liquid culture, cells may be grown in L-broth and antibiotic through about 10 or more doublings. The phage are harvested by standard procedures (see Sambrook et al., (1989) Molecular Cloning, 2nd ed. (1989), supra, for preparation of M13 phage) as further modified as described below. Growth in liquid culture may be more convenient because of the size

of the libraries, while growth on solid media likely provides less chance of bias during the amplification process.

For affinity enrichment of desired clones, about 10^3 to 10^4 library equivalents (a library equivalent is one of each recombinant -- 10^4 equivalents of a library of 10^9 members is $10^9 \times 10^4 = 10^{13}$ phage) are incubated with hapten (ligand) to which the desired protein (e.g., antibody) is sought. The hapten is in one of several forms appropriate for affinity enrichment schemes. In one example the hapten is immobilized on a surface or particle, usually anchored by a tether of enough length (3 to 12 carbons, for example) to hold the hapten far enough away from the surface to permit free interaction with the antibody combining site. The library of phage bearing antibodies is then panned on the immobilized hapten generally according to the procedure described in the Example section below.

A second example of hapten presentation is hapten attached to a recognizable ligand (again with a tether of some length). A specific example of such a ligand is biotin. The hapten, so modified, is incubated with the library of phage and binding occurs with both reactants in solution. The resulting complexes are then bound to streptavidin (or avidin) through the biotin moiety. The streptavidin may be immobilized on a surface such as a plastic plate or on particles, in which case the complexes (phage-antibody-hapten-biotin-streptavidin) are physically retained; or the streptavidin may be labelled, with a fluorophore, for example, to tag the active phage/antibody for detection and/or isolation by sorting procedures, e.g., on a fluorescence - activated cell sorter.

The phage bearing antibodies without the desired specificity are removed by washing. The degree and stringency of washing required will be determined for each protein of interest. A certain degree of control can be exerted over the binding characteristics of the antibodies recovered by adjusting the conditions of the binding incubation and the subsequent washing. The temperature, pH, ionic strength, divalent cations concentration, and the volume and duration of the washing will select for antibodies within particular ranges of affinity for the hapten. Selection based on slow dissociation rate, which is usually predictive of high affinity, is the most practical route. This may be done either by continued incubation in the presence of a saturating amount of free hapten, or by increasing the volume, number, and length of the washes. In each case, the rebinding of dissociated antibody-phage is prevented, and with increasing time, antibody-phage of higher and higher affinity are recovered.

Additional modifications of the binding and washing procedures may be applied to find antibodies with special characteristics. The affinities of some antibodies are dependent on ionic strength or cation concentration. This is a useful characteristic for antibodies to be used in affinity purification of various proteins when gentle conditions for removing the protein from the antibody are required. Specific examples are antibodies which depend on Ca^{++} for binding activity and which released their haptens in the presence of EGTA. See Hopp et al., Biotechnology 6:1204-1210 (1988). Such antibodies may be identified in the recombinant antibody library by a double screening technique isolating first those that bind hapten in the presence of Ca^{++}, and by subsequently identifying those in this group that fail to bind in the presence of EGTA.

Antibodies with certain catalytic activities may be enriched in groups of antibodies with high affinity for reactants (substrates and intermediates) but low affinity for products. A double screen to enrich for antibodies with these characteristics may be useful in finding antibodies to catalyze certain reactions. Further, catalytic antibodies capable of certain cleavage reactions may also be selected. One category of such reactions is the cleavage of a specific end group from a molecule. For example, a catalytic antibody to cleave a specific amino acid from an end of a peptide may be selected by immobilizing the peptide and panning the antibody library under conditions expected to promote binding but not cleavage (e.g., low temperature, particular ionic strength, pH, cation concentration, etc., depending on the nature of the end group and the cleavage reaction) and followed by a wash. This allows antibodies that recognize the end group to bind and become immobilized, and from this group will come those capable of cleavage. To find those capable of cleavage, the conditions are shifted to those favorable for cleavage. This step will release those antibody-phage capable of cleaving themselves free of the immobilized peptide.

An alternative way to accomplish this is to pan for antibodies that bind to the specific end group by attaching that end group to a bond different from that to be cleaved (a non-peptide bond, for example). By subsequent panning (of the positive phage from the first screen) on the end group attached via the proper bond under cleavage conditions, the non-binding fraction will be enriched for those with the desired catalytic activity.

To elute the active antibody-phage from the immobilized hapten, after washing at the appropriate stringency, the bound (active) phage are generally recovered by eluting with pH shift. For example, pH2 or pH11 may be used, which is then neutralized and the eluted phage are amplified by infecting or transforming the host cells. Examples of such hosts are E. coli, MC1061-F'KAN or K91. The cells are then grown as tetracycline resistant colonies. The colonies are scraped up and the extruded phage are purified by standard procedures as before. These phage are then used in another round of affinity enrichment (panning), and this cycle is repeated until the desired level of enrichment is reached or until the target phage are no longer enriched relative to the background phage. Repeated rounds of panning and intervening amplifications can provide levels of enrichment exceeding 10^7-fold. To isolate individual clones, phage from the final round of panning and elution are infected into cells or their DNA is transformed into cells and grown on agar (usually L-agar) and antibiotics (usually tet) to form well separated individual colonies, each of which is a clone carrying phage genomes with both V_H and V_L sequences. The single stranded DNA from phage particles extruded from each

colony may be isolated and DNA coding for the V_H and V_L fragments sequenced. The replicative form of the phage DNA (double stranded) may be isolated by standard means and the DNA in the cloning sites (V_H and V_L sequences) recloned into a vector designed for gene product expression in prokaryotes or eukaryotes to obtain larger amounts of the particular antibodies selected in the screening process.

Phage identified as having an antibody recognized by the target ligand are propagated as appropriate for the particular phage vector used. For fd-tet this is done in a liquid culture of rich medium (L-broth, for example) with antibiotic (Tet) selection. The phage are harvested and DNA prepared and sequenced by standard methods to determine the DNA and amino acid sequence of the particular antibody.

The DNA may be recloned in a suitable eukaryotic or prokaryotic expression vector and transfected into an appropriate host for production of large amounts of protein. Antibody is purified from the expression system using standard procedures. The binding affinity of the antibody is confirmed by well known immunoassays with the target antigen or catalytic activity as described in Harlow and Lane, Antibodies. A Laboratory Manual, Cold Spring Harbor, N.Y. (1988).

The following example is offered by way of illustration, not by way of limitation.

EXAMPLE I

This Example describes a procedure for isolating an antibody which binds to a preselected antigen, and thereby isolating the nucleotide sequences which encode the antibody. An immunoglobulin expression library is prepared from mouse spleen cells using a filamentous phage, fd. The antibody heavy chains are expressed as fusion proteins with coat protein pIII. Phage particles which contain the antibody of the desired binding specificity are isolated by means of a panning procedure using the preselected antigen.

Construction of Vectors

A filamentous bacteriophage vector, fdTetBI, was constructed from the tetracycline resistance transducing vector fdTet (Zacher et al., 1980). The cloning site in fdTetB1 was engineered into the N-terminal region of gene III, and comprises two non-complementary BstXI sites separated by 18 bases. This was accomplished by first removing a BstXI restriction site that was already present in the TN10 region of fdTet. RF DNA was digested with BstXI restriction endonuclease, followed by the addition of T4 polymerase to remove the protruding 3' termini. Blunt ended molecules were then ligated and electrotransformed into MC1061. RF DNA was isolated from several tetracycline resistant transformants and was digested with BstXI restriction endonuclease. A clone which was not digested with this enzyme was selected for insertion of the cloning site by site directed mutagenesis (Kunkel et al., Meth. Enzymol., 154:367-382 (1987)) with the oligonucleotide S'TATGAGGTTTTGCCAGACAACTGGAACAGTTTCAGCGGA GTGCCAGTAGAAT-GGAACAACTAAAGG-3'. Insertion of the correct mutagenic sequence was confirmed by dideoxy sequencing of RF DNA that was isolated from several tetracycline resistant transformants.

Construction of Degenerate Oligonucleotide Library

Vector fdTetBSN was constructed to contain a pair of non-complementary SfiI sites in the TN10 region. This site is the ultimate destination of the expression cassette from the light chain library. fdTetBI was opened at the unique HindIII site. Two synthetic, complementary oligonucleotides were kinased and annealed to form the structure

```
AGCTGGCCGCAGCGGCCGCGGCCGCGGCCGGTCCGGCC
    CCGGCGTCGCCGGCGCCGGCGCCGGCCAGGCCGGTCGA
```

This oligo was ligated to the HindIII site of fdTetB1 by standard methods. A properly ligated and circularized molecule should contain no HindIII sites, so after inactivating the ligase, the reaction was recut with HindIII to linearize those plasmids that had not taken up an insert. This material was then transformed into *E. coli* and selected on tetracycline. The structure of the resulting molecules was verified and these were designated fdTetBSN.

To construct the vector fdTetSXNS, fdTetBSN was opened at the BstXI sites and an oligonucleotide, annealed to give the following structure

```
    GCC TCG AGA GCA CGA CGT ACT AGT GCT TGT
A GAT CGG AGC TCT CGT GCT GCA TGA TCA CG
```

was ligated to the BstXI sites of fdBSN to provide unique XhoI and SpeI sites to receive the heavy chain fragments.

The plasmid vector pVL was constructed by digesting pUC19 with A1wNI and polishing the ends to bluntness. An SfiI site was introduced by ligating the annealed oligonucleotides

GGCCGCAGCGGCC

CCGGCGTCGCCGG

to the blunted AlwNI site.

The AatII site of this modified pUC19 was opened and a fragment containing the light chain expression cassette described by Huse et al., flanked on the downstream side (with respect to transcription) by an SfiI site of the sequence GGCCGGTCCGGCC. This plasmid now contains (in clockwise direction) SfiI site, B-lactamase gene, a lac Z promoter, a ribosome binding site, a pel B signal peptide sequence, SacI site, XbaI site, translation termination site(s), and SfiI site.

Construction of the separate heavy and light chain libraries

cDNA sequences representing the antigen binding domains of the heavy and light antibody chains are synthesized from the RNA of antibody producing cells in the manner described by Huse et al., supra. Spleen cells are used from mice that have been immunized with the preselected antigen to which the desired antibody binds. The cDNA is amplified by PCR using primers annealing to the regions common to many antibody sequences that flank the variable antigen binding domains. The primers also contain appropriate restriction site sequences as described below. The amplified, double-stranded fragments are digested with the appropriate restriction nucleases and ligated to the compatible sites in the respective vectors. For the heavy chain fragments this is accomplished by the incorporation of XhoI sites into the 5'-PCR primers and SpeI sites into the 3'-primers. These sites are then exposed by digestion and ligated to the corresponding sites in the digested fdTetSXNS vector. The ligation products were transformed by electroporation into E. coli MC1061 and, after outgrowth, selected on tetracycline.

For the light chain library, the light chain fragments are amplified with 5'-PCR primers containing SacI sites and 3'-primers containing XbaI sites. The sites on the fragments are opened and ligated to the corresponding sites in pVL. The ligation products are then transformed by electroporation into E. coli MC1061. After outgrowth, transformants were recovered by selection on ampicillin.

Construction of the combined heavy-light chain expression library

DNA from the light chain library (in pVL) was digested with SfiI. The larger of the two resulting fragments, which contains the light chain expression cassette, is isolated and ligated in several-fold molar excess to DNA of the heavy chain library (in fdTetSXNS) that has been digested with SfiI. The ligation products are transformed by electroporation into E. coli MC1061 cells. After a period of outgrowth, the transformants are double-selected on ampicillin (100 ug/ml) and tetracycline (20 ug/ml). The resulting transformants constitute combined library of heavy and light chain fragments in fdTetSXNS.

The methods of cDNA synthesis, restriction digestions, fragment phosphorylation, and ligation are essentially as prescribed in Sambrook et al., supra. In all cases, DNA to be transformed is ethanol precipitated in the presence of 0.3M sodium acetate and resuspended in water. Electrotransformations are done as described by Dower et al., supra. After 1 hour of non-selective outgrowth at 37°C in SOC medium (2% Bacto tryptone, 0.5% Bacto yeast extract, 10mM NaCl, 2.5mM KCl, 10mM MgCl2, 10mM MgS04, 20mM glucose) and an aliquot is removed and several dilutions are plated on LB plates containing the appropriate antibiotic (20 ug/ml tetracycline, 100 ug/ml ampicillin, or both). The remainder of the transformation is used to inoculate one liter of L-broth containing the appropriate antibiotic and is grown for various times to amplify the library. Phage are isolated from this culture and stored at 4°C.

Isolation of Phage

Purified phage from liquid cultures are obtained by clearing the supernatant two times by centrifugation, and precipitating phage particles with polyethylene glycol (final concentration 3.3% polyethylene glycol 8000, 0.4M NaCl). Following centrifugation, phage pellets are resuspended in TBS (50 mM Tris-HCl, pH 7.5, 150mM NaCl) and stored at 4° C. Phage are also isolated from plate stocks in this manner, after scraping colonies from the agar surface and resuspending in L-broth.

Affinity Purification

Approximately 10^3 - 10^4 library equivalents of phage are reacted overnight with 1 µg purified antibody at 4° C. The

mixture is panned by a procedure as follows. A 60 x 15 mm polystyrene petri plate is coated with 1 ml of streptavidin solution (1 mg/ml in 0.1 M $NaHCO_3$, pH 8.6, 0.02% NaN_3) and is incubated overnight at 4° C. The following day the streptavidin solution is removed. The plate is filled with 10 ml blocking solution (30 mg/ml BSA, 3 µg/ml streptavidin in 0.1 M $NaHCO_3$, pH 9.2, 0.02% NaN_3) and incubated for 2 hours at room temperature. Two micrograms of biotinylated goat anti-mouse IgG (BRL) are added to the antibody-reacted phage library and incubated for 2 hours at 4°C. Immediately before panning, blocking solution is removed from streptavidin coated plate, and the plate is washed 3 times with TBS/0.05% Tween® 20. The antibody-reacted phage library is then added to the plate and incubated for 30 minutes at room temperature. Streptavidin coated agarose beads (BRL) may also be used for this affinity purification. The phage solution is removed and the plate is washed ten times with TBS/0.05% Tween® 20 over a period of 60 minutes. Bound phage are removed by adding elution buffer (1 mg/ml BSA, 0.1 N HCl, pH adjusted to 2.2 with glycine) to the petri plate and incubating for 10 minutes to dissociate the immune complexes. The eluate is removed, neutralized with 2M Tris (pH unadjusted) and used to infect log phase F'-containing bacterial cells. These cells are then plated on LB agar plates containing tetracycline (20 µg/ml), and grown overnight at 37°C. Phage are isolated from these plates as described and the affinity purification process was repeated for two to three rounds. After the final round of purification, a portion of the eluate is used to infect cells and plated at low density on LB tetracycline plates. Individual colonies are transferred to culture tubes containing 2 ml LB tetracycline and grown to saturation. Phage DNA is isolated using a method designed for the Beckman Biomek® Workstation (Mardis and Roe., Biotechniques, 7:840-850 (1989)) which employs 96-well microtiter plates. Single stranded DNA is sequenced by the dideoxy method using Sequenase® (U. S. Biochemicals) and an oligonucleotide sequencing primer (5'-CGATCTAAAGTTTTGTCGTCT-3') which is complementary to the sequence located 40 nucleotides 3' of the second BstXI site in fdTetBI.

It is evident from the above that compositions and methods are provided which substantially increase the ability to isolate nucleotide sequences which encode proteins of interest, particularly antibodies, from a large DNA library. This is especially encouraging, in that these methods and compositions may be employed to recover or produce de novo many proteins, and particularly monoclonal antibodies, useful as therapeutic or prophylactic compositions, diagnostic reagents, catalytic compounds, etc. previously obtainable only by extensive experimentation, if at all. The invention herein provides a means to circumvent many of the difficulties associated with traditional methods of monoclonal antibody technology, and particularly human monoclonal antibody technology. Further, the proteins identified by the present invention are produced by recombinant means, providing additional advantages of, inter alia, convenience, substantial purity and economics.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.

Claims

1. The use of a bacteriophage to display a multichain protein, wherein a first chain of the multichain protein is fused to a coat peptide on the outer surface of the bacteriophage, and a second chain of the multichain protein is complexed with the first chain.

2. A method for screening a DNA library for nucleotide sequences which encode a multichain protein comprising first and second polypeptide chains, which multichain protein binds specifically to a ligand, comprising:

   effecting bacteriophage expression vector transformation of a host cell with:-

      (i) a first nucleotide sequence member of the library that encodes the first chain fused to a sequence encoding a coat peptide of the bacteriophage; and

      (ii) a second nucleotide sequence member of the library that encodes the second chain fused to a sequence encoding a signal peptide that directs periplasmic secretion of said second chain;

   cultivating the transformed cell under conditions suitable for expression and assembly of bacteriophage particles and the multichain protein, wherein the multichain protein is displayed on the outer surface of the bacteriophage particles, optionally wherein expression of DNA library sequence members is inducible, induction of expression of the DNA library sequences preferably being delayed until assembly of at least one complete bacteriophage particle has occurred;

   selecting bacteriophage particles encoding the multichain protein by means of the ligand and, if desired, further

comprising the step of isolating the nucleotide sequences which encode the first and second chains of the multichain protein from the selected bacteriophage particles; optionally wherein the bacteriophage are harvested from the host cell culture before the selecting step.

3. A method according to claim 2, wherein the bacteriophage expression vector is a filamentous bacteriophage.

4. A method according to claim 3, wherein the filamentous phage is M13, fd or f1, preferably fd or a derivative thereof.

5. A method according to any one of claims 2 to 4, wherein the coat peptide is a pIII protein.

6. A method according to any one of claims 2 to 5, wherein the signal sequence is an Omp A, pel B, phoA, pIII or β-lactamase signal sequence.

7. A method according to any one of claims 2 to 6, wherein the transformed host cells are lysed after cultivation and the bacteriophage particles are selected from cellular debris; optionally wherein the bacteriophage particles encoding the protein of interest are enriched by repeating the selection step at least once.

8. A method according to any one of claims 2 to 7, wherein the multichain protein comprises an antibody or a binding fragment thereof.

9. A method according to claim 8, wherein the multichain protein is a Fab fragment.

10. A method of claim 8, wherein said first nucleotide sequence member of the library encodes a protein which comprises an antibody heavy chain variable region, optionally said antibody heavy chain variable region being located at the amino terminus of said coat protein on said bacteriophage surface.

11. A method according to any one of claims 2 to 10, wherein the first and second DNA library members comprise amplified cDNA.

12. A method as claimed in any one of claims 2 to 11 further comprising the step of incorporating the multichain protein into a therapeutic, prophylactic or diagnostic composition.

Patentansprüche

1. Die Verwendung eines Bakteriophagen, um ein Multikettenprotein darzustellen, worin eine erste Kette des Multikettenproteins fusioniert ist zu einem Hüll-Peptid der äußeren Oberfläche des Bakteriophagen und eine zweite Kette des Multikettenproteins einen Komplex bildet mit der ersten Kette.

2. Ein Verfahren zum Screening einer DNA-Genbank für Nucleotidsequenzen, die ein Multikettenprotein kodiert, umfassend eine erste und zweite Polypeptidkette, die das Multikettenprotein spezifisch an einen Liganden binden, umfassend:

Bewirkung einer Bakteriophagen-Expressionsvektor-Transformation einer Wirtszelle mit

(i) einer ersten Nucleotidsequenz der Genbank, die die erste Kette, fusioniert an eine Sequenz, kodierend ein Hüll-Peptid des Bakteriophagen, kodiert und

(ii) eine zweite Nucleotidsequenz der Genbank, die die zweite Kette, fusioniert an eine Sequenz, kodierend ein Signalpeptid, das die periplasmatische Sekretion der zweiten Kette bewirkt, kodiert;

Kultivierung der transformierten Zelle unter Bedingungen, die geeignet zur Expression und zum Zusammenbau der Bakteriophagen-Partikel und des Multikettenproteins sind, wobei das Multikettenprotein dargestellt wird auf der äußeren Oberfläche der Bakteriophagen-Partikel, gegebenenfalls wobei die Expression der DNA-Genbank-Sequenzen induziert wird, wobei die Induktion der Expression der DNA-Genbank-Sequenzen gegebenenfalls verzögert wird bis zum Zusammenbau von mindestens einem vollständigen Bakteriophagen-Partikel;

Selektion der Bakteriophagen-Partikel, die das Multikettenprotein kodieren mit Hilfe des Liganden und, falls gewünscht, ferner umfassend den Schritt der Isolierung der Nucleotidsequenzen, die die erste und zweite Kette des Multikettenproteins der ausgewählten Bakteriophagen-Partikel kodieren; gegebenenfalls, wobei der Bakteriophage von der Wirtszellkultur vor dem Selektionsschritt geerntet wird.

3. Verfahren gemäß Anspruch 2, wobei der Bakteriophagen-Expressionsvektor ein filamentöser Bakteriophage ist.

4. Verfahren gemäß Anspruch 3, wobei der filamentöse Phage M13, fd oder fl, bevorzugt fd oder ein Derivat davon ist.

5. Verfahren gemäß einem der Ansprüche 2 bis 4, wobei das Hüll-Peptid ein pIII-Protein ist.

6. Verfahren gemäß einem der Ansprüche 2 bis 5, wobei die Signalsequenz eine Omp A, pel B, phoA, pIII oder β-Lactamase-Signalsequenz ist.

7. Verfahren gemäß einem der Ansprüche 2 bis 6, wobei die transformierten Wirtszellen nach der Kultivierung lysiert werden und die Bakteriophagen-Partikel von den Zelltrümmem getrennt werden, gegebenenfalls wobei die Bakteriophagen-Partikel, die das interessierende Protein kodieren, angereichert werden durch mindestens einmalige Wiederholung des Selektionsschrittes.

8. Verfahren gemäß einem der Ansprüche 2 bis 7, wobei das Multikettenprotein einen Antikörper oder ein Bindungsfragment hiervon umfaßt.

9. Verfahren gemäß Anspruch 8, wobei das Multikettenprotein ein Fab-Fragment ist.

10. Verfahren gemäß Anspruch 8, wobei die erste Nucleotidsequenz der Genbank ein Protein kodiert, das die variable Region einer schweren Kette eines Antikörpers umfaßt, gegebenenfalls ist die variable Region der schweren Kette des Antikörpers an dem Amino-Terminus des Hüll-Proteins auf der Bakteriophagen-Oberfläche lokalisiert.

11. Verfahren gemäß einem der Ansprüche 2 bis 10, wobei die ersten und zweiten DNA-Genbank-Mitglieder amplifizierte cDNA umfassen.

12. Verfahren, wie in einem der Ansprüche 2 bis 11 beansprucht, ferner umfassend den Schritt der Inkorporation des Multikettenproteins in eine therapeutische, prophylaktische oder diagnostische Zusammensetzung.

Revendications

1. Utilisation d'un bactériophage pour former une protéine multichaîne, dans laquelle une première chaîne de la protéine multichaîne est fusionnée à un peptide de l'enveloppe situé sur la surface externe du bactériophage, et une seconde chaîne de la protéine multichaîne est complexée avec la première chaîne.

2. Procédé de criblage d'une banque d'ADN pour obtenir des séquences nucléotidiques codant pour une protéine multichaîne comprenant une première et une seconde chaînes polypeptidiques, ladite protéine multichaîne se liant spécifiquement à un ligand, ledit procédé comprenant les étapes consistant à :

   transformer une cellule hôte avec un bactériophage comme vecteur d'expression comprenant :

   (i) une première séquence nucléotidique membre de la banque, codant pour la première chaîne fusionnée à une séquence codant pour un peptide de l'enveloppe du bactériophage ; et
   (ii) une seconde séquence nucléotidique membre de la banque, codant pour la seconde chaîne fusionnée à une séquence codant pour un peptide signal qui dirige la sécrétion périplasmique de ladite seconde chaîne ;

   cultiver la cellule transformée dans des conditions adaptées pour l'expression et l'assemblage des particules de bactériophage et de la protéine multichaîne, la protéine multichaîne étant formée sur la surface externe des particules de bactériophage, l'expression des séquences membres de la banque d'ADN étant éventuellement inductible, l'induction et l'expression des séquences de la banque d'ADN étant de préférence retardées jusqu'à assemblage d'au moins une particule complète de bactériophage ;

sélectionner les particules de bactériophage codant pour la protéine multichaîne au moyen du ligand et, si nécessaire, isoler ultérieurement les séquences nucléotidiques qui codent pour les première et seconde chaînes de la protéine multichaîne à partir des particules de bactériophage sélectionnées ;
recueillir éventuellement le bactériophage à partir de la culture de cellule hôte avant l'étape de sélection.

3. Procédé selon la revendication 2, dans lequel le bactériophage vecteur d'expression est un bactériophage filamenteux.

4. Procédé selon la revendication 3, dans lequel le phage filamenteux est le phage M13, fd ou f1, de préférence fd ou un dérivé de celui-ci.

5. Procédé selon l'une quelconque des revendications 2 à 4, dans lequel le peptide d'enveloppe est la protéine pIII.

6. Procédé selon l'une quelconque des revendications 2 à 5, dans lequel la séquence signal est une séquence signal pour Omp A, pel B, phoA, pIII ou β-lactamase.

7. Procédé selon l'une quelconque des revendications 2 à 6, dans lequel les cellules hôtes transformées sont lysées après culture et les particules de bactériophage sont sélectionnées à partir des débris cellulaires, les particules de bactériophage codant pour la protéine d'intérêt étant éventuellement enrichies par répétition de l'étape de sélection au moins une fois.

8. Procédé selon l'une quelconque des revendications 2 à 7, dans lequel la protéine multichaîne comprend un anticorps ou fragment de liaison de celui-ci.

9. Procédé selon la revendication 8, dans lequel la protéine multichaîne est un fragment Fab.

10. Procédé selon la revendication 8, dans lequel ladite première séquence nucléotidique membre de la banque code pour une protéine qui comprend une région variable de chaîne lourde d'anticorps, ladite région variable de chaîne lourde d'anticorps étant éventuellement localisée à l'extrémité amino-terminale de ladite protéine d'enveloppe sur la surface dudit bactériophage.

11. Procédé selon l'une quelconque des revendications 2 à 10, dans lequel les premier et second membres de la banque d'ADN comprennent de l'ADNc amplifié.

12. Procédé selon l'une quelconque des revendications 2 à 11, comprenant en outre une étape d'incorporation de la protéine multichaîne dans une composition thérapeutique, prophylactique ou diagnostique.